**Black Magazines as a Mirror to 1960-70s America: *A machine learning perspective***

**Kushal K Dey** [**kkdey@uchicago.edu**](mailto:kkdey@uchicago.edu) **Department of Statistics, University of Chicago**

**Lei Sun** [**sunl@uchicago.edu**](mailto:sunl@uchicago.edu) **Department of Statistics, University of Chicago**

**Lauren M. Jackson** [**lmjackson@uchicago.edu**](mailto:lmjackson@uchicago.edu) **Department of English, University of Chicago**

**Abstract**

In many ways, the turbulent times of the 1960s and early 1970s America, though fraught with social unrest, foreign crisis, and cultural revolution, played a great role in shaping a generation. In addition to the many significant developments of the time, this era was also a witness to changing African American society, lifestyle and culture. This evolution was reflected in African American magazines and publications of the time.

In this paper, we use machine learning and natural language processing tools to analyze two leading African American magazines of the time – Negro Digest (which later changed its name to Black World) and Ebony – for a period of 16 years from 1961 to 1976. We observe primarily one major

shift in media narrative during this time period that coincides with the end of the civil rights movement in 1968. Narratives connecting African Americans to race, discrimination, protests and movements, which are more pronounced during the civil rights era gradually make way for narratives related to African American art, music and culture post 1968. We also show how the contextual relatedness of words changes across time and how the gradual shift from “negro” to “black” as the more popular identity term impacts the media narratives.

**Introduction**

Since the latter half of the 20th century, news media has served to not only reflect and represent society’s morale, but must be understood as an active participant in the social and political climate. As the times change, media narratives change and vice-versa. These narratives are not simply determined as matter of two-party politics and a wider, more involved consideration of the discourse that shapes our society is required to feel out their contours and index evolutions in public opinion. Besides the usual social and political influences, one issue of key interest is the evolution of *blackness* as a staple idiom with African American or Black American publications.

In an attempt to address this issue, we focus on a time period in American history that has had monumental influence in shaping up Black American society and culture of today – the 1960s and early 1970s. This was a period of great social and political turmoil. The freedom struggle of the Black Americans seeking equal rights and social justice gained nationwide attention – a movement popularly known as the civil rights movement. With the assassination of the movement’s greatest pioneer Dr. Martin Luther King in 1968, this movement eventually entered a new phase with the rise of the Black Power and the Black Arts movements. This period was also marked by foreign crisis in the form of the Vietnam war (1961-1975) and the many anti-war movements it spawned and the rise of a new counterculture and the hippie revolution in late 1960s and early 1970s. A detailed timeline of the events of the time period of our study is provided in Supplementary Figure 1.

In this paper, we attempt to understand how the social and political movements and developments of this era shaped the media narratives through the lens of two leading monthly African American publications of the time, The *Negro Digest* - which changed its name to *Black World*, and *Ebony*. However subjectively comparing and contrasting these magazines and their changing narratives over time is no easy task. To this end, we employ state-of-the-art machine learning and natural language processing tools to objectively quantify the differences in narratives between these magazines and also for the same magazine portal over time.

Our analysis point sto one major shift in media narrative happening around the time of Dr. King’s assassination in 1968 that marked the end of the civil rights movement. The narrative during the civil rights movement era (1961-1968) appears to be driven more by social and political issues focusing on racism, discrimination, protest movement and freedom struggle. The post 1968 narrative however is driven by narratives related to the Black arts movement and also narratives that talk about social and economic development. Also, during this era, we see a gradual decline in the identity word “negro” which is largely substituted in later years by the identity word “black”. This shift in identity reveals interesting examples of narratives that are contextually more (less) similar to one identity term compared to the other, or are not affected by this shift at all.

**Methods and Materials**

Founded originally in 1942 by the Chicago based publisher John Johnson, the *Negro Digest* magazine was revived in 1961 and was consistently in circulation for a period of 16 years from 1961 to 1976. In May 1970, this magazine was re-named *Black World*.

The *Ebony* magazine, also founded by John Johnson, was founded in 1945 and has been in continuous circulation since then up to this day. Both these magazines have been archived around 1960 onwards in Google Books. We extracted the text data from these archived versions of the magazines, for the time frame of 16 years (1961-1976) when both *Negro Digest/Black World* and *Ebony* were in circulation.

For each magazine, we first downloaded its Google Book version in Portable Document Format (.pdf) and then used Optical Character Recognition (OCR) to extract the text from these files into a .txt format. This was followed by manual inspection to test for the OCR quality, which was found to be reasonably high. Next, we converted all the text to lowercase letters and removed any special characters and numbers from the extracted text. Some of the bigrams, say *civil rights*, were then converted into single word (*civil\_rights*) so that we can demonstrate how this phrase as a whole shapes the narrative instead of its component words (*civil* and *rights*). Other examples of phrases that were converted into single words were *martin luther king, dr king, black power, black panther, malcolm x*. We aggregate the monthly processed text data into a single text data for each year. This processed yearly text data is ultimately used as input in all our statistical modeling frameworks.

The statistical methods applied for defining narratives and identifying narrative shifts can be categorized into two groups – ones that are based on word frequency changes over time and the others that account for changes in contextual similarity and co-occurrence patterns of different words over time.

*Word Frequency Analysis*

For this approach, we first e compile the frequencies of occurrence of each word in the yearly text data of each magazine. This can be processed into a matrix of counts *C = ((cng))* for each magazine, with the rows representing the year of publication, the columns representing the words in the vocabulary and each entry *cng*  denoting the number of times word *g* occurs in the yearly text of year (*1960+n)*.

Next we employ an unsupervised clustering framework that can identify narrative groups purely based on the word frequency matrix *C* for each magazine. The model we propose here is popularly known as topic models in natural language processing (Blei et al). This model fits a multinomial model to each row of the matrix *C.*

, , ……., ) ~ Mult (, , , ……., )

where

= with and

Here K is the number of topics or *frequency narratives* and

is the relative contribution of word g to the topic k, whereas

is the proportional contribution or grade of membership of the topic k in the yearly text for year (1960+n). The estimated

are plotted in using a stacked bar chart, also called a STRUCTURE plot (Pritchard). The estimated values can used to identify the top words that distinguish one topic from the rest

using the method proposed in Dey et al 2017 (*ExtractTopFeatures()* in R package **CountClust**) [ref].

*Word Context Analysis*

The word frequency approach does not account for the semantic structure of the sentences and the contextual relationship between words. For example, words like *music* and *songs* are contextually close as they will likely have similar words as neighbors where they occur in the text. A machine learning approach that can quantify this contextual relationship between each pair of words is *word2vec* (ref). As the name suggests, this method projects each word to a vector in a D dimensional space

(D user specified) using a neural network modeling framework. The projection mechanism entails if two words are contextually close, the angular distance between their projections would be proportionately close and hence the cosine of the angular distance can be used as a measure of the contextual similarity of the word pairs.

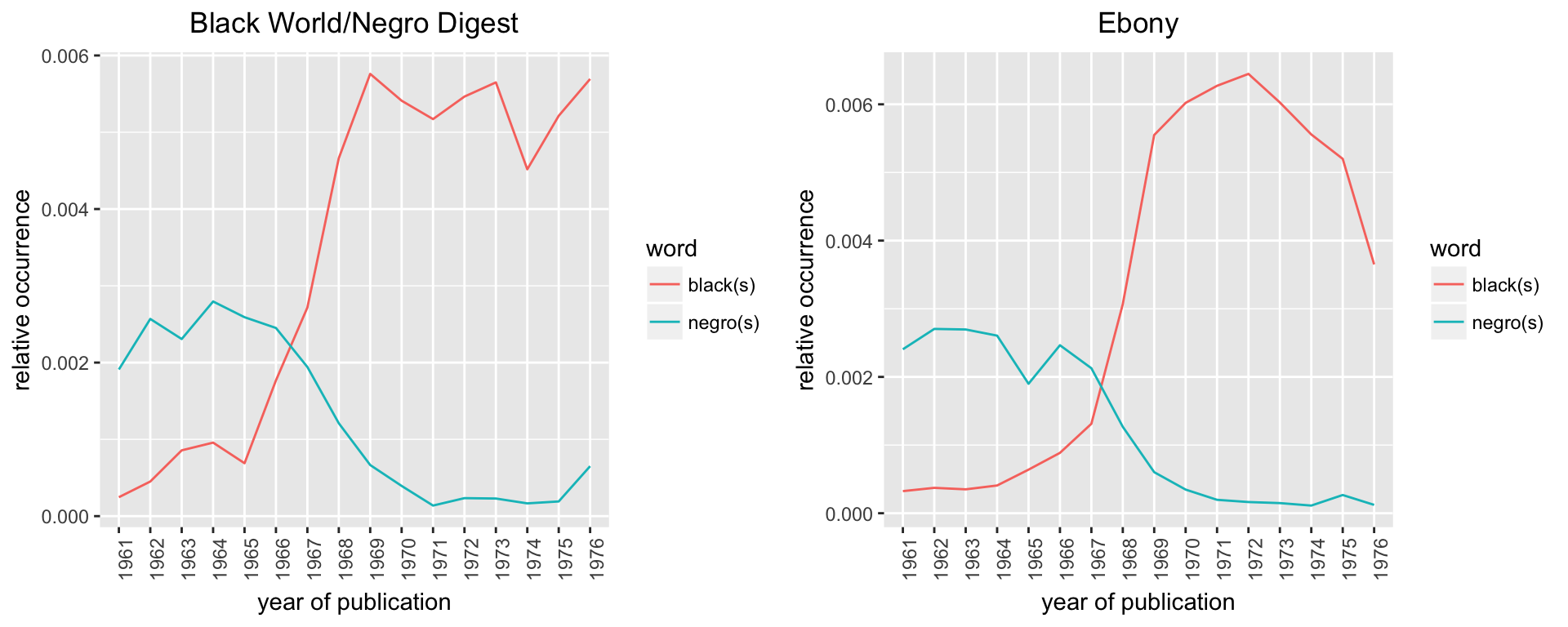
Here we fit the word2vec model on the text data for each year of a magazine to estimate the contextual (cosine) similarity of word pairs for that specific year. Then for each pair of words or word sets, we investigate the trends of contextual similarity over time.

This approach lets us identify word pairs whose contextual similarity increases or decreases or stays at the same level over time.

For our application, we used the Continuous Bag of Words (CBOW) version of the word2vec model. The dimension of projection space was chosen to be 100, and the context for each word was defined by 10 words immediately around it in the processed text (without special characters and numbers). Words that occur less than or equal to 3 times in the yearly text were removed from consideration.

**Results**

Figure 1 shows the trends in relative frequencies of the identity words “black(s)” and “negro(s)” in the two magazines. We observe that “negro” as an identity term gradually falls out of favor while “black” becomes the more pronounced identity word in the post civil-rights era – a shift that has been extensively studied already (ref).



**Figure 1**: Yearly trends in relative occurrences of the words black(s) and negro(s) for the two magazines – Black World/Negro Digest and Ebony. The identity “negro” appears to fall out of favor as “black” becomes the more popular identity word.

We applied the topic model to the counts matrix C for each magazine separately . To remove the clear identity bias in the narratives as evident in Figure 1, words like black(s), negro(s), afro-american etc were removed from the analysis. Figure 2 shows the results from topic model fits with K=2 for the two magazines. In both the magazines, the pre-1968 (civil-rights era) issues show high membership in one topic/narrative whereas the post 1968 issues show high membership in the other, topic with the time period from 1966-1968 marking the transition from one narrative to the other.